DOCUMENT RESUME

ED 457 184		TM 033 280
AUTHOR	Kogos, Stephen C., Jr.	
TITLE	Repeated Measures Designs and the Sphericity Assumption.	
PUB DATE	2000-00-00	
NOTE	12p.	
PUB TYPE	Reports - Evaluative (142)	
EDRS PRICE	MF01/PC01 Plus Postage.	
DESCRIPTORS	*Counseling; Multivariate Analysis; *Psychology; *Research Design	
IDENTIFIERS	*Repeated Measures Design; *Sphericity Tests	

ABSTRACT

This paper presents an overview of repeated measures designs, designs that measure one group over time, also called within-subjects designs. Both univariate and multivariate designs are discussed, and advantages and disadvantages of each of these designs are considered. The sphericity assumption is examined, and methods are presented to correct violations of this assumption. The sphericity assumption is met if all the differences between pairs of treatment condition scores are equally variable, and if variances of differences for all treatment conditions are homogenous. The paper also offers some suggestions for incorporating these statistical procedures into research in the field of counseling psychology. (SLD)

RUNNING HEAD: REPEATED MEASURES

Repeated Measures Designs and the Sphericity Assumption

Stephen C. Kogos, Jr.

University of Southern Mississippi

Abstract

The current paper presents an overview of repeated measures designs. Both univariate and multivariate designs are discussed. Advantages and disadvantages of each of these designs are also considered. Next, the sphericity assumption is examined, and methods to correct violations of this assumption are presented. Finally, the author offers some suggestions for incorporating these statistical procedures into research in the Counseling Psychology field.

Repeated Measures Designs and the Sphericity Assumption

Introduction

Repeated measures designs are those that measure one group over time. Within the group, each subject is given two or more treatments. Another name for this type of test that may be more familiar to consumers is the within subjects design. The simplest example of this design is the correlated t-test, which most researchers and graduate students are familiar with. In the correlated t-test, the same subjects are pretested and posttested (more than one measurement) on a dependent variable with an intervening treatment. To use an example pertinent to counseling psychology, a researcher may want to determine the efficacy of relaxation training on clients' anxiety levels. To do this, the researcher would measure subjects' anxiety using an instrument such as the Beck Anxiety Inventory (pretesting). Then, relaxation training might be done for four counseling sessions. Finally, subjects would take the Beck Anxiety Inventory again (posttesting).

The univariate repeated measures design is recommended in several research scenarios. First, whenever random assignment of subjects is not possible, the univariate repeated measures design is frequently used. In situations when a subject pool is limited, random assignment may not be feasible, so the researcher may want to consider using this design. Wells (1998) states, "random assignment is not always possible or even the most desirable experimental design; repeated measures designs can prove more advantageous" (p. 3). Next, repeated measures designs are recommended with research that focuses on performance trends over time. Utilizing this design allows the researcher to examine the effect of an intervention at different levels or different points in time and note any trends in the effects. The researcher may want to look for linear, quadratic, or cubic relationships and graph these. Finally, repeated measures designs are

advantageous in drug studies, in which a researcher wants to assess the most effective level or dose of a drug. To use another example related to counseling psychology, perhaps a researcher believes that a new drug will serve as an effective anti-depressant, but is uncertain about the most effective dosage. The researcher could give four subjects each 0 mg, then a Beck Depression Inventory (BDI), 50 mg, followed by a BDI; then 100 mg, and a BDI; and finally, 150 mg., and a BDI. Such a design would have several advantages over a traditional between-subjects one-way ANOVA. (see SPSS printouts attached). These advantages will now be discussed.

The repeated measures design offers several distinct advantages over other research designs. First of all, the repeated measures design is economical as regards the number of subjects required for a given experiment (Tanguma, 1999). Because the same subjects take part in all conditions of the study, fewer are required. This is particularly beneficial when the subject pool is limited. Although a liability of this design is that it may be more time consuming, Roberts (1998) argues that the results from the repeated measures analysis have greater power benefits for the analysis. Another important advantage of a repeated measures design is the control of subject variability, or individual differences. Lewis (1993), in Wells (1998), stated that selecting only one group of subjects and having that group participate in all the treatment conditions seems to guarantee that any observed differences among the treatment conditions is due strictly to the effects of the treatment. In other words, in a traditional between subjects design that does not employ repeated measures, even though random assignment may be used, it is always possible that an outlier may influence the results of the study. A final advantage of this design may be the reduction of error variance. The hallmark of the within-subjects design is that it attempts to minimize error variance due to individual variation by having each subject serve as his or her own control (Heppner et al., 1992). Basically, because the subjects remain the same, the variance

due to subjects can be partitioned out of the error variance term, which makes any statistical tests more sensitive, or powerful (Stevens, 1996, cited in Wells, 1998).

Although the cited advantages make the repeated measures design an appealing choice, there are certain disadvantages to this design that researchers should be wary of before selecting this design. One of these is carryover or practice effects. Practice effects can be either positive (such as improvement) or negative (ex. boredom or fatigue). It is important to realize that within the same study, even if some subjects show improvement and some show a decline in performance, practice effects may have occurred as they can be both positive and negative. A second disadvantage is order effects, which refers to the order in which subjects receive the treatments. Researchers can account for order effects through a technique known as counterbalancing. This is a way of ordering treatments so that each treatment is administered an equal number of times first, second, third, and so on, in particular sequences of conditions given to different subjects (Keppel & Zedeck, 1989, in Tanguma, 1999). Tanguma (1999) proposes that two situations can arise when using counterbalancing. These are: (1) there is an even number of levels of the treatment conditions within a way or factor; or (2) there is an odd number of levels of treatment conditions. These two scenarios must be considered separately.

Girden (1992), cited in Tanguma (1999) and Wells (1998) stated that when the number of levels of the treatment conditions, k, is an even number and the number of subjects, n, is some multiple of k, researchers should use the following pattern for assigning subjects to treatments: 1, 2, n, 3, n-1, 4, n-2, etc. In this formula, each number refers to a treatment level Each subsequent order is derived by adding "1" to the numbers of the preceding order. An example of counterbalancing for four subjects is provided in Table 1. As Tanguma notes, "it follows that

Table1	Order of Treatments			
Subject	1	2	3	4
1	A	B	D	C
2	B	C	A	D
3	C	D	B	A
4	D	A	C	B

each treatment precedes each of the other treatments exactly once" (p. 236). In other words, A precedes B, C, and D exactly once; B precedes each of A,C, and D exactly once; and D precedes A,B, and C exactly once.

When there is an odd number of levels of treatment conditions, the first order of presentation is derived the exact way as it was with an even number of levels. However, the rest of the levels are obtained by reversing the order of the first order, and then repeating the procedure. This is illustrated in Table 2.

Table 2. Counterbalancing for Five Subjects

	Order of Treatments				
Subject	1	2	3	4	5
1	A	B	E	C	D
2	D	C	E	B	A
3	E	D	A	C	B
4	A	E	B	D	C
5	B	A	C	E	D

Finally, Kieffer, 1998 (in Wells, 1998) reminds researchers that although counterbalancing is a necessary condition in conducting repeated measures designs, it is not sufficient by itself. It does not rectify all problems that occur in these analyses.

A final disadvantage that can be problematic is latency effects. This is a delayed effect of a treatment that does not become evident until a second treatment has already begun. Research suggests that this may be even more difficult to control for than the practice or carryover effects.

In fact, there may be no means of "controlling" this once it occurs. Perhaps the best way to defend against this problem is to take preventative measures so that it does not occur. One suggestion might be to allow for sufficient time between treatments; however, this decision must be weighed in regards to how much time an experimenter has to conduct a study. Overall, this is a decision left up to the researcher, who should be aware of this potential problem.

Assumptions of Univariate Repeated Measures Analyses

Stevens (1996), cited in Tanguma (1999) provided three assumptions of the univariate repeated measures analysis. These are: (a) independence of observations; (b) multivariate normality; and (c) sphericity (also known as circularity). Of these assumptions, sphericity may be the most complex, so the remainder of discussion of assumptions will focus on that assumption. The sphericity assumption is viewed as a necessary and sufficient condition for employing a univariate repeated measures analysis (Huynh & Feldt, 1970, in Wells, 1998). This assumption is met if all the differences between pairs of treatment condition scores are equally variable and if variances of differences for all treatment conditions are homogenous (Girden, 1992, in Wells, 1998). In other words, people should respond similarly across treatments. If subjects do not respond similarly across treatments, then variability due to subjects has occurred. Girden recommended this formula to determine sphericity:

where is the variance of one set of scores under treatment A, is the variance of another set of scores under treatment B, and is the covariance of the two sets of scores. Basically, the researcher should find the differences in the treatment scores, calculate the variances, and then see if they are different.

Correcting Violations of this Assumption

A common solution to a violation of the sphericity assumption is a correction factor known as epsilon. Epsilon measures nonsphericity. The upper limit of epsilon is 1.0, in which case sphericity is perfectly met. The lower limit of epsilon is 1/k-1, where k = # of treatment conditions in the study. Reductions in epsilon, then, indicate increasing degrees of nonsphericity. The epsilon correction factor is multiplied times the degrees of freedom to obtain a corrected degrees of freedom. There are two epsilon correction factors that are commonly used. These are: 1) the Geisser-Greenhouse, and 2) the Huynh-Feldt. The Geisser-Greenhouse produces a very conservative estimate for the degrees of freedom. The Huynh-Feldt epsilon correction index is not as conservative; it produces an overestimation of the true epsilon value. This may result in too small of an F critical value, which may lead to Type I errors. Keiffer (1998), in Wells, 1998 explained that because these two correction indices may produce different results, many researchers suggest averaging the two indices. However, it seems that when forced to choose one or the other, the researcher would be advised to choose the Geisser-Greenhouse as it is a more conservative estimate.

Multivariate Repeated Measures Analysis

The multivariate approach treats the repeated measurements as separate dependent variables generated by one individual. The most common approach to defining the new dependent variable set is to transform the K dependent variables into k-1 linearly independent pairwise difference scores, and then perform the analysis on these new dependent variables (Minke, 1997, in Wells, 1998). In sum, the researcher will not use the original scores anymore; rather, the differences in the scores will be utilized. One major advantage of this approach is that the measurements can be correlated with one another (i.e. the sphericity assumption becomes unnecessary). Related to this advantage, another may be that this assumption may approach the

researcher's reality more closely, as the measurements can now be correlated with each other. There are also disadvantages to the multivariate approach though. One of these is that statistical significance becomes more difficult to attain. This is because the researcher has now lost the advantage of repeatedly measuring participants because now each measurement is a separate dependent variable. Another potential problem is that sample size becomes more important than in the univariate approach.

Univariate or Multivariate?

Despite much debate among researchers, no clear consensus has been reached on this issue. It seems that univariate may be more desirable if the sphericity assumption holds (which is rare) because the univariate design would be more powerful. However, when the assumption is violated, the situation can become very complicated. The multivariate approach is probably less powerful as n decreases, and more powerful when n increases. To further elaborate on this, Stevens (1996) suggests that in order to employ the multivariate design, the researcher should have at least ten more subjects than the number of measures. Overall, it seems that making this decision may depend largely on the data and the degree to which the sphericity assumption is violated.

Having introduced the repeated measures designs and the sphericity assumption, a final purpose of this paper is to consider how this research design might be incorporated into the field of counseling psychology. In counseling psychology, researchers are often interested in measuring participants' responses on more than one occasion. Some more obvious examples of repeated measures contexts include process research, interactional or discourse research, developmental studies, outcome research, profile analysis, and single-subject studies (Ellis, 1999).

From a practitioner's perspective, what seems obvious is that what most clients seek in coming to therapy is some form of change. These changes can be cognitive, emotional, behavioral, or any combination of these. It seems pertinent, then, that a client and clinician are able to collaborate to measure whether as therapy progresses, some form of change is in fact occurring. Repeated measures designs are the logical choice for making these type of assessments. This procedure may be as simple as having all clients who present for depression complete the brief, 21-item Beck Depression Inventory (BDI) before the first session of therapy. Then, as the client continues in therapy, he or she would be required to periodically complete the BDI.

For those interested in a health psychology/behavioral medicine approach, possibilities abound for using the repeated measures design. Frequently, measuring change in this setting may involve recording the number of cigarettes smoked in a week, the number of pounds lost in a month, the amount of time spent exercising in a week, or other measurements of the frequency of a specific behavior. These type of presenting problems lend themselves perfectly to the repeated measurements design.

Overall, the potential for using repeated measures designs in counseling psychology research is endless. Researchers seeking to minimize the number of subject required in their study and seeking a powerful design should certainly consider utilizing this design in future studies.

References

Ellis, M.V. (1999). Repeated measures designs. The Counseling Psychologist, 27,(4), 552-578.

Heppner, P.P., Kivlighan, D.M., & Wampold, B.E. (1992). Research design in counseling. Pacific Grove, CA: Brooks/Cole Publishing Company.

Roberts, J.K. (1998). Repeated measures ANOVA and regression. Paper submitted to Dr. Bruce Thompson in partial fulfillment of Final Comps, College Station, TX.

Stevens, J. (1996). Applied multivariate statistics for the social sciences: Third edition. Mahwah, NJ: Lawrence Erlbuam Associates, Inc.

Tanguma, J. (1999). Analyzing repeated measures designs using univariate and multivariate methods: A primer. Advances in Social Science Methodology, 5, 233-250.

Wells, R.S. (1998). Conducting repeated measures analyses using regression: The general linear model lives. Paper presented at the annual meeting of the Mid-South Educational Research Association, New Orleans, LA.

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse: University of Maryland
ERIC Clearinghouse on Assessment and Evaluation
1129 Shriver Laboratory
College Park, MD 20742
Attn: Acquisitions

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
1100 West Street, 2nd Floor
Laurel, Maryland 20707-3598

Telephone: 301-497-4080
Toll Free: 800-799-3742
FAX: 301-953-0263
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com